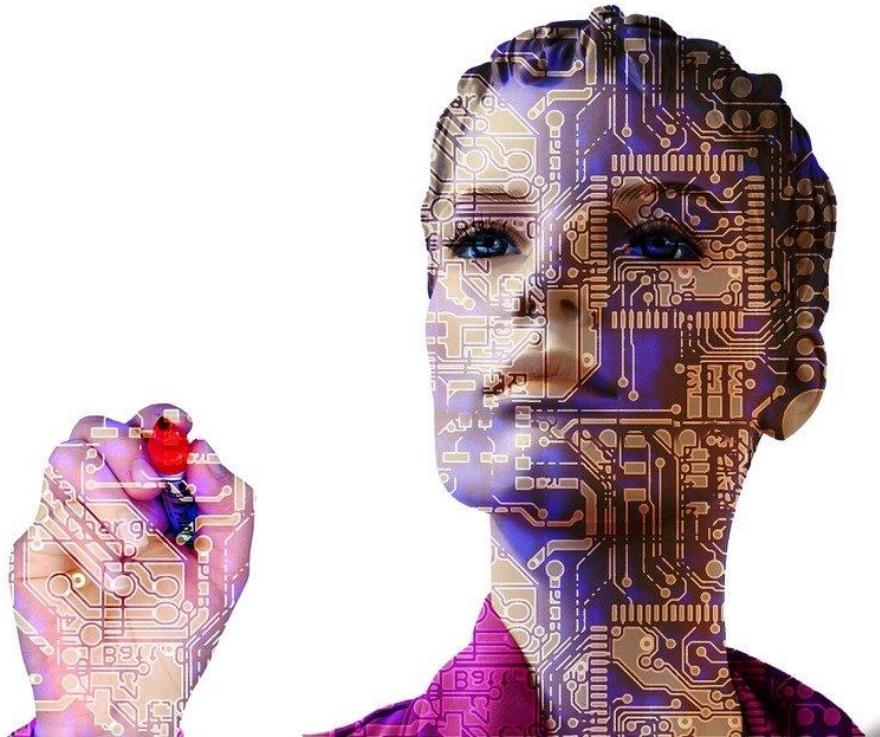


Guía de uso de la Inteligencia Artificial en el Sector Público



MONTAÑA MERCHÁN ARRIBAS

CUERPO SUSPÉRIOR DE SISTEMAS Y
TECNOLOGÍAS DE LA INFORMACIÓN

Con la colaboración de Carmen Barrera

2020

TÍTULO: Guía de uso de la inteligencia artificial en el sector público 2020

Elaboración y coordinación de contenidos: Montaña Merchán Arribas

Octubre de 2020

Esta publicación está disponible en
<http://digitalrevolution.info>

Colección: Tecnologías emergentes



El presente documento está bajo la licencia Creative Commons Reconocimiento-No comercial-Compartir Igual versión 4.0 España

Usted es libre de: **copiar** y redistribuir el material en cualquier medio o formato y **adaptar**, remezclar, transformar y crear a partir del material.

Bajo los siguientes términos:

- Atribución — Usted debe dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciatante.
- NoComercial — Usted no puede hacer uso del material con propósitos comerciales.
- Compartirlgual — Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la misma licencia del original..

Esto es un resumen legible por humanos del texto legal (la licencia completa) disponible en <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

Contenido

Introducción	4
Que es la IA?	5
Un repaso a la tecnología	6
Machine learning – Def	6
Red neuronal-Def	7
Robótica –Def	8
El procesamiento del lenguaje natural (PLN)-Def	9
Principios de la IA	10
Oportunidades de la IA para el sector público.....	13
Big data analytics y calidad del dato	14
Machine learning.....	16
Asistentes inteligentes chatbots	16
Procesamiento de Lenguajes Naturales	17
Procesamiento de imágenes	18
Robótica inteligente (cognitiva)	18
Capacitación	19
Metodología para aplicar la IA en un proyecto.....	21
Factores para aplicar la IA en un proyecto.....	22
Fases del proyecto IA	23
Impacto	29
Seguridad.....	29
Responsabilidad	30
Ética	31
Transparencia	31
Objetividad.....	32
Protección de datos y regulación.....	32
Anexo 1 Glosario inicial de la IA	33

Introducción

La Inteligencia Artificial (IA) es uno de los pilares de la llamada ciencia cognitiva, esta ciencia busca entender la forma en que funciona la mente humana.

La expresión *inteligencia artificial* se acuñó por primera vez en la conferencia de Dartmouth en 1956, pero el concepto de que una máquina haga tareas sola es idea de Turing en 1936. El inicio de la IA puede considerarse en **Deep Blue**¹ un ordenador de IBM capaz de jugar al ajedrez y ganar a Gari Kaspárov. Aunque tuvo su éxito, el verdadero auge de la IA resurge actualmente gracias a la capacidad de gestionar grandes volúmenes de datos. Ahora tenemos una mayor potencia computacional que cada vez es más accesible y barata con la que podemos tratar una inmensa cantidad de datos que se generan diariamente. Esta sinergia ha provocado un renacimiento de la IA que parece más prometedor que los anteriores. En ella se están viendo significativos avances en todos los campos de la IA, siendo tal vez el de los sistemas de aprendizaje automático (Machine Learning-ML), los que más han aprovechado esta nueva situación.

La IA forma parte de nuestro modelo de vida, en los asistentes domésticos; los traductores de idioma; la generación automática de subtítulos; los sistemas de reconocimiento de patrones; los sistemas de diagnóstico de enfermedades; los coches sin conductor; los sistemas anti-spam y los sistemas de detección avanzada de amenazas usados en Ciberseguridad. Por citar sólo algunas aplicaciones conocidas.

Consideramos que la Inteligencia Artificial (IA) tiene el potencial de transformar la economía y la sociedad de una forma tan profunda, que provocará cambios sociales, económicos y laborales.

El objetivo de esta Guía de uso de la IA en la AP, **GUIAP**, es el de proponer una orientación a las Administraciones Públicas Españolas, para reconocer el valor que puede aportar la IA a sus funciones.

Se ha estructurado en 4 bloques:

1. una somera descripción de las tecnologías que engloba la IA,
2. las oportunidades que ofrece para el sector público y
3. una metodología para abordar un proyecto de IA,
4. el impacto de la IA en la administración

¹ Deep Blue fue una supercomputadora desarrollada por el fabricante estadounidense IBM para jugar al ajedrez de la década de los 90.

Que es la IA?

No hay una definición universalmente aceptada de la Inteligencia artificial. La Inteligencia artificial es el nombre genérico que engloba varias tecnologías y que dependen también del campo científico al que se aplican (biología, medicina, agricultura, etc). Quizás sea la Universidad de Stanford la que ha logrado una de las definiciones más claras:

“la IA es una ciencia y un conjunto de técnicas que están inspiradas, aunque típicamente operan de una manera diferente, en cómo los seres humanos utilizan su sistema nervioso y su cuerpo para sentir, aprender, razonar y actuar”.

Por tanto, la IA es una combinación de tecnologías relacionadas con la representación del conocimiento, los sistemas de razonamiento y el aprendizaje automático, que conjuntamente proporcionan la capacidad de entender y aprender, adaptándose a distintas circunstancias, proporcionando hallazgos (nuevos conocimientos), tomando decisiones o exhibiendo nuevos comportamientos.

Los expertos en ciencias de la computación Stuart Russell y Peter Norvig diferencian varios tipos de inteligencia artificial:

- A. Sistema que realizan tareas de forma similar a como lo hacen las personas (robots)
- B. Sistema que automatizan actividades como la toma de decisiones, la resolución de problemas y el aprendizaje
- C. Sistemas que emulan o imitan el pensamiento lógico racional de los humanos (sistemas expertos, agentes inteligentes).

Aunque la meta a largo plazo es tener una IA de propósito general, el estado actual de la investigación y la tecnología nos ubica en una “Narrow AI” limitada a tareas y metas concretas. Dentro de IA se incluyen campos como el procesamiento de lenguaje natural, visión computacional, reconocimiento del habla, las redes neuronales, sistemas basados en el conocimiento o planificación automática y el desarrollo de sistemas autónomos.

En todos estos campos y tecnologías se lleva trabajando durante mucho tiempo, pero en los últimos años se están viendo significativos avances, en especial en el área de aprendizaje automático (Machine Learning - ML) y el software de Robotic Process Automation RPA , que es donde vamos a centrar esta guía.

Un repaso a la tecnología

La Inteligencia Artificial se puede dividir de la siguiente forma

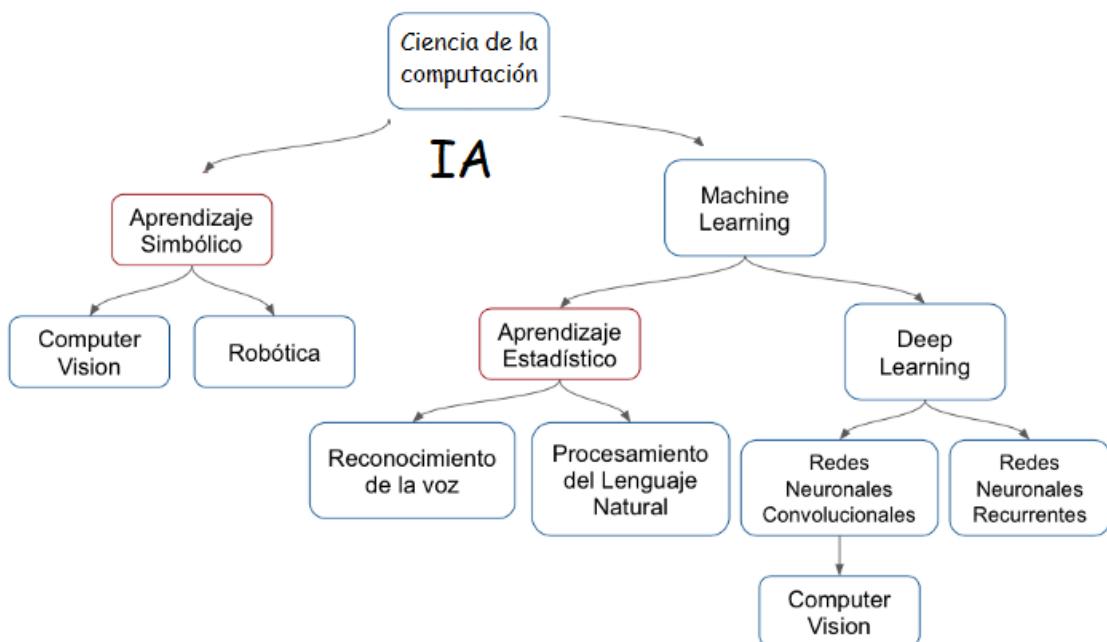


Ilustración 1 clasificación de IA

Machine learning – Def

En el código de la programación tradicional se establecen las reglas necesarias para procesar los datos de entrada y obtener unos resultados. Sin embargo, en Machine learning (ML) se trata de resolver un problema en el que las reglas no son tan claras; el sistema recibe datos de entrada y las respuestas esperadas, y su objetivo es obtener las reglas [Ilustración 1 Una vez obtenidas, estas reglas se aplicarán a nuevos datos para producir respuestas originales.

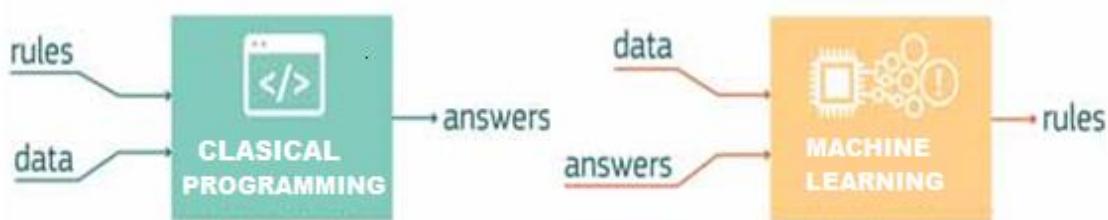


Ilustración 2. Fuente: Chollet, F. 2017. Deep Learning and Python. Manning

Un sistema de ML necesita ser entrenado, proporcionándole un conjunto de datos suficientemente grande y representativo del problema. Este entrenamiento suele ser un proceso interactivo e iterativo, hasta que el modelo obtenido es satisfactorio y puede explotarse, habitualmente incorporándose como un componente más del sistema.

ML se desarrolla a partir de algoritmos y de datos necesario a para entrenar los algoritmos

Algoritmos de ML	Problemas que resuelve
Aprendizaje Supervisado permite aprender a partir de datos de entrenamiento etiquetados con las respuestas. Es común que ese etiquetado lo realice un humano	Problemas de Clasificación : Predice la categoría, dentro de un conjunto finito predeterminado, a la que pertenece el nuevo dato. Problemas de Regresión : Consisten en predecir el valor de una variable numérica en función de los valores que haya tomado antes. Es un problema con un espacio de soluciones continuo, dado que cualquier valor es permitido
Aprendizaje No-Supervisado permite aprender desde los datos, para obtener nuevo conocimiento de éstos buscando los patrones comunes; no necesita conocimiento previo sobre qué buscar en los datos. Este tipo de aprendizaje es útil cuando existen enormes cantidades de datos no estructurados, y no se tiene un análisis de estos que permita su clasificación previa.	Problemas de agrupación o “clustering” . Consisten en encontrar factores comunes en los datos para agruparlos en grupos o clústeres que no son aparentes a la inspección humana. Diagnósticos
Aprendizaje Semi-Supervisado combina los dos anteriores, a partir de un conjunto de casos etiquetados, utiliza técnicas de aprendizaje no supervisado para etiquetar casos de forma masiva	Problemas de detección de anomalías La clasificación automática de documentos
Aprendizaje de Refuerzo : el sistema realiza acciones y el aprendizaje se realiza a partir del refuerzo positivo o negativo recibido de las mismas.	Procesos de prueba y error para evaluar nuevas estrategias Juegos y simulaciones Sistema de control de un vehículo

En los últimos años ha surgido una nueva línea de ML que está teniendo mucho éxito, el aprendizaje profundo (Deep Learning – DL) y que aborda estos mismos problemas mediante redes neuronales complejas.

Red neuronal-Def

Una **red neuronal**, inspirada en el cerebro humano, está compuesta por una red de pequeñas unidades de procesamiento (análogamente a las neuronas) con muchas conexiones ponderadas entre ellas. Durante la etapa de entrenamiento, recibe como entrada los datos ejemplo, y se van ajustando los pesos de las conexiones para

conseguir obtener la salida deseada. Es necesaria una fase adicional para comprobar el comportamiento de la red neuronal sobre nuevos ejemplos y verificar que la tarea se ha aprendido correctamente.

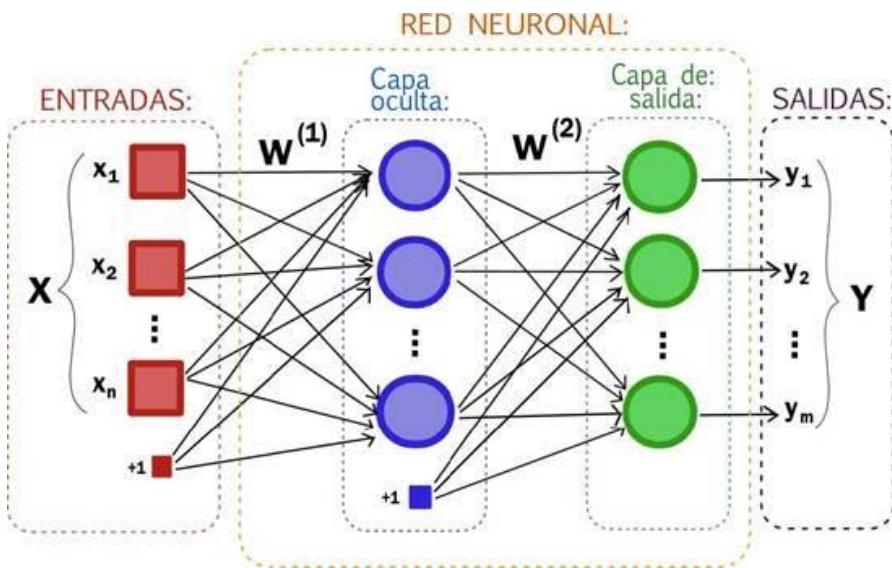


Ilustración 3: Red neural

Aunque las redes neuronales no son nuevas, el planteamiento actual de DL presenta modelos con múltiples niveles de neuronas entre la entrada y la salida que permiten aprender la relación entre la entrada y salida en pasos sucesivos. Con ello se consigue un enfoque más **preciso** y con menos necesidad de guía humana. Este modelo DL exige un alto grado de complejidad; el elevado número de neuronas aumenta exponencialmente las conexiones, y el ajuste de los pesos requiere una capacidad de cálculo muy grande así como algoritmos matemáticos y estadísticos muy sofisticados. Para conseguir esa precisión se requieren una gran cantidad de datos de entrenamiento y que estos sean representativos del problema.

Robótica -Def

La robótica es la ciencia que se encarga del diseño, fabricación y empleo de máquinas automáticas y programables cuya finalidad es realizar tareas repetitivas y otras actividades. Una de las áreas que engloba es el software de *Robotic Process Automation* o RPA. La Automatización Robótica de Procesos busca procesar automáticamente tareas repetitivas. RPA utiliza robots de software para realizar acciones como “scraping”, agregar datos, limpiarlos y comunicarse con otras aplicaciones y personas para ejecutar trabajos repetitivos.

Una condición indispensable para el uso de cualquier RPA es su aplicación en procesos estructurados, con inputs y outputs basados en datos existentes en el mundo “digital”. Los RPA no “aprenden” de su experiencia y no están capacitados para resolver las excepciones que encuentren al ejecutar sus tareas.

El procesamiento del lenguaje natural (PLN)-Def

El procesamiento del lenguaje natural (PLN) es la “capacidad de las máquinas para comprender e interpretar el lenguaje humano de la forma en que está escrito o hablado.

La combinación de tecnologías aumenta la potencia de resolver problemas. Por ejemplo, el software de Automatización Robótica de Procesos (RPA) puede extraer automáticamente el contenido de un contrato que se adjunta a un email que llega a la bandeja de entrada de una firma de abogados. A continuación, el RPA puede trasladar el contenido del contrato a una herramienta de categorización, que extraerá datos complejos, como las partes en el contrato, los términos de una cláusula particular o los afectados por un procedimiento legal. La herramienta de procesamiento del lenguaje puede detectar e identificar sus cláusulas y otras partes relevantes, tales como el Título, las Partes, la Fecha, el Término, la asignación, el Cambio de control, la Auditoría, la Ley aplicable, la Fuerza mayor, la Indemnización, la Limitación de responsabilidad, etc

Cada una de estas tecnologías constituye verdaderas especialidades dentro del campo de la IA, donde cada una de ellas posee características propias y se ocupan de resolver problemas de la vida real.

En este sentido y ante un cierto problema, el especialista debe identificar la o las tecnologías que mejor se adapten para ese problema; lo que en la actualidad se erige en uno de los principales cuellos de botella para cada una de estas tecnologías, y para la IA en general.

En el apartado x se dan pautas para asociar tecnologías a problemas.

Principios de la IA

Cualquier lista de principios que las organizaciones han presentado sobre el uso de la inteligencia artificial se resume en una frase:

La inteligencia artificial debe anteponer a la gente y al planeta.

Todas las discusiones sobre la IA ética giran alrededor de que el uso de la IA sea compatible con la viabilidad de la supervivencia humana. Tanto la Unión Europea, como la OCDE o los gobiernos de los distintos países están definiendo políticas para evitar los riesgos que puede suponer el uso de la IA, en cuanto a la libertad, la responsabilidad, la igualdad, los derechos humanos y los valores democráticos.

El grupo de expertos de la Comisión Europea publicó un primer borrador de las directrices éticas en diciembre de 2018. En ella se describen **7 directrices** para guiar a las empresas hacia **sistemas de Inteligencia Artificial confiables**.

En abril de 2019 la Comisión Europea publicó la *comunicación: Building Trust in Human Centric Artificial Intelligence* en el que se postulan que para lograr una "IA confiable", se necesitan tres componentes: (1) debe cumplir con la ley, (2) debe cumplir con los principios éticos y (3) debe cumplir debe ser robusto.

Sobre la base de estos tres componentes y los valores europeos establecidos en ese documento², se identifican **siete requisitos** clave que las aplicaciones de IA deben respetar para ser consideradas confiable.

Supervisión humana:	Los sistemas de AI deben permitir sociedades equitativas apoyando los derechos fundamentales de los humanos, y no disminuir, limitar o desviar la autonomía de estos. Hay que garantizar que las autoridades públicas tengan la capacidad de ejercer sus competencias de supervisión conforme a sus mandatos
Solidez técnica y seguridad	La inteligencia artificial requiere que los algoritmos sean lo suficientemente seguros, fiables y sólidos como para enfrentar errores o inconsistencias durante todas las fases del ciclo de vida de los sistemas de inteligencia artificial.
Privacidad y control de los datos	Los ciudadanos deben tener control total sobre sus propios datos, además estos datos no se utilizarán para perjudicar a los ciudadanos o discriminarlos. Deben garantizarse la privacidad y la protección de datos en todas las fases del ciclo vital del sistema de IA

² The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights.

Transparencia	Se debe garantizar la trazabilidad de los sistemas de IA. En la medida de lo posible el proceso de toma de decisiones algorítmico, debe ser explicable.
Diversidad, no discriminación y equidad	Los sistemas de AI deben considerar toda la gama de habilidades y requisitos humanos, y garantizar la accesibilidad.
Bienestar social y ambiental	Velar porque los datos no tengan sesgos, que sean completos. Los sistemas de inteligencia artificial deben utilizarse para mejorar el cambio social positivo y mejorar la sostenibilidad y la responsabilidad ecológica.
Responsabilidad	Deben establecerse mecanismos para garantizar la responsabilidad de los implementadores en el funcionamiento de los sistemas y en sus resultados. La posibilidad de auditar los sistemas de IA es fundamental. Los potenciales impactos negativos de los sistemas de IA deben señalarse, evaluarse, documentarse y reducirse al mínimo.

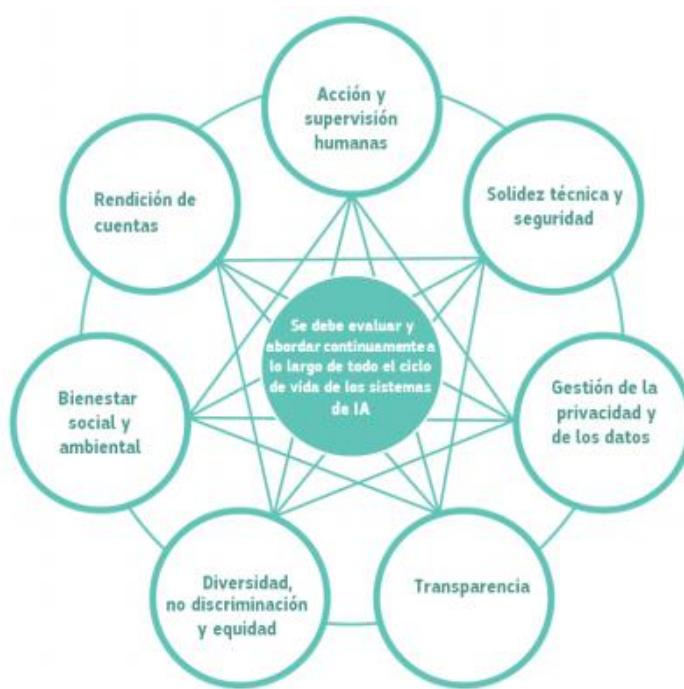
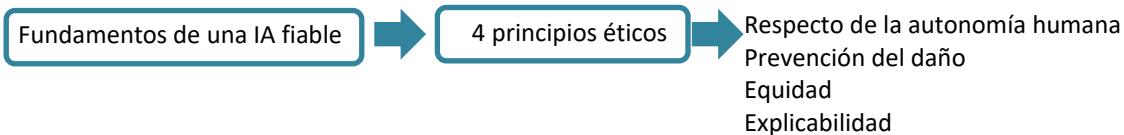


Ilustración 4 Interrelaciones existentes entre los siete requisitos: todos tienen idéntica importancia, se apoyan entre sí

Para cumplir los requisitos anteriormente expuestos, cabe utilizar tanto métodos técnicos como de otro tipo. Dichos métodos abarcan todas las fases del ciclo de vida de un sistema de IA. Es importante tener en cuenta que **debería llevarse a cabo una evaluación constante** de los métodos empleados para cumplir los requisitos; asimismo, se debería informar y justificar en todo momento los cambios introducidos en el proceso de aplicación de estos. Dado que los sistemas de IA evolucionan sin cesar y actúan en un entorno dinámico, la realización de la IA fiable es un proceso continuo, como muestra la ilustración

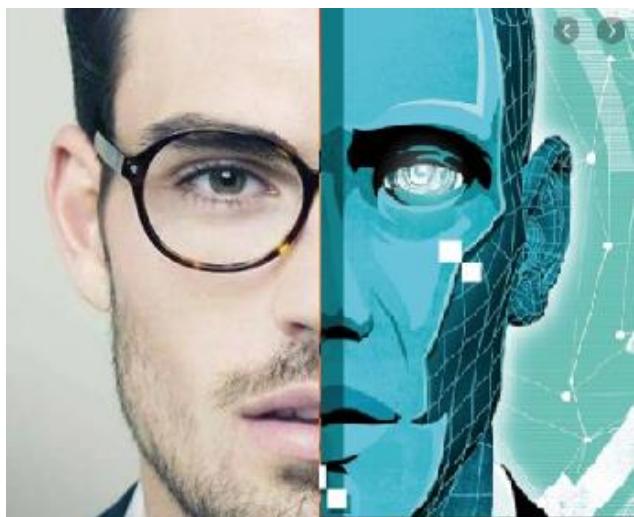
La Unión Europea también ha publicado el Libro Blanco sobre la IA (19/02/2020) para lograr que sea “fiable, basada en la excelencia y la confianza. El objetivo es ofrecer alternativas para facilitar un desarrollo de la inteligencia artificial seguro y fiable en Europa, que respete plenamente los valores y los derechos de los ciudadanos de la UE. Los pilares fundamentales del presente Libro Blanco son: la excelencia y la confianza.

Como resumen para conseguir una IA confiable, 4 principios éticos:



Respeto de la autonomía humana	La IA debería ayudar a las personas en el entorno laboral y aspirar a crear empleos útiles y garantizar el respeto de la libertad y la autonomía de los seres humanos.
Prevención del daño	Los sistemas de IA no deberían provocar daños ni perjudicar de cualquier otro modo a los seres humanos ni al entorno natural o a todos los seres vivos. Al contrario puede ser una herramienta poderosa para mejorar las oportunidades para las personas con discapacidad. Por tanto se espera que su funcionamiento sea confiable y seguro dentro de ciertos parámetros.
Equidad	El desarrollo de sistemas de IA debe fomentar la igualdad de oportunidades.
Explicabilidad	Esta característica es crucial para conseguir que los usuarios confíen en los sistemas de IA y para mantener dicha confianza. No sólo es necesario identificar a la entidad responsable de la decisión sino que hay que poder explicar los criterios por los que se toman decisiones que afectan a los ciudadanos. Es cierto que en los algoritmos de <i>caja negra</i> no siempre resulta posible saber cómo se ha generado un resultado, pero en ese caso hay que adoptar otras medidas de trazabilidad, auditoria o transparencia para asegurar la explicabilidad.

Oportunidades de la IA para el sector público.



Las tecnologías que se enmarcan en la IA no son de propósito general, no solucionan cualquier problema, sino que se aplican a tareas específicas. La IA puede funcionar muy bien para resolver problemas de reconocimiento de patrones o de la voz, toma de decisiones, sistemas de predicciones, ... pero las soluciones no son intercambiables.

La aplicabilidad de la inteligencia artificial depende en gran parte de la

disponibilidad de datos de calidad. Ya se utilice para crear algoritmos de IA o para elaborar políticas públicas, los datos son fundamentales para entender los problemas

Las máquinas solo poseen una pequeña porción de nuestras ricas y versátiles habilidades de aprendizaje

Los sistemas de inteligencia artificial no son capaces de identificar causa-efecto ni la causalidad ni pueden inferir el contexto, a diferencia de la inteligencia natural.

¿En qué se puede usar la IA en el sector público?: la IA es buena para resolver problemas específicos y bien definidos. Puede proporcionar múltiples beneficios en:

- La automatización de tareas repetitivas de los procedimientos que consumen tiempo, liberando a los empleados públicos de esas tareas (tecnología softbots RPA).
- La mejora de las decisiones públicas proporcionando información más precisa, pronósticos y predicciones que conduzcan a mejores resultados (aprendizaje automático).
- La mejora de los servicios públicos mediante el uso de IA para proporcionar soluciones más cercanas y personalizadas (chatbots)
- La simulación de sistemas complejos que permiten experimentar con diferentes opciones y detectar consecuencias no deseadas antes de tomar decisiones.
- El reconocimiento de imágenes para agricultura, seguridad pública, etc (Redes Neuronales Artificiales)

LA IA ES MUY BUENA PARA RESOLVER PROBLEMAS ESPECÍFICOS Y BIEN DEFINIDOS, pero no es de propósito general. Necesita un volumen de datos para funcionar bien y no infiere contexto por lo que los modelos necesitan ser monitorizados y revisados.

Algunas de las dimensiones en las que la IA está produciendo avances espectaculares en el ámbito de la Administración Pública son:

- **La salud:** múltiples aplicaciones de la IA en el ámbito de la salud se están utilizando ya, como diagnóstico automático por imagen, cirugía robótica, predicción de riesgos de enfermedad, enfermeras virtuales, control de las prescripciones de medicamentos, etc.
- **La industria:** ya se habla de la revolución de la Industria 4.0, que permite aprovechar al máximo las aplicaciones de IA y RI en seguridad y productividad.
- **La educación:** la IA permitirá adaptar la enseñanza a las capacidades y necesidades de cada alumno
- **La defensa y seguridad:** las aplicaciones de la IA y RI en la defensa y seguridad son múltiples y serán tratadas con detalle en este trabajo.
- **La agricultura:** el reconocimiento de imágenes y el aprendizaje automático también se está utilizando en la planificación de cultivos.
- **El aprendizaje:** La realidad aumentada y la realidad virtual incrementarán el aprendizaje remoto. La educación es y será más personalizada. La IA están abriendo una diversidad de posibilidades en el entrenamiento y la formación.

Algunas soluciones se adaptarán mejor que otras a la Administración. Lo que es indiscutible es que para aplicar IA se van a necesitar un cierto volumen de datos y datos de calidad y es por lo que no podemos obviar en esta Guía el BIG DATA y la analítica de datos.

A continuación se ofrece una lista de tecnologías y cómo se está aplicando en la Administración Pública.

Big data analytics y calidad del dato

Big Data es un término que describe un gran volumen de datos, tanto estructurados como no estructurados. **Big data analytics** es el proceso de utilizar los algoritmos de análisis de datos para extraer información no trivial a partir de grandes volúmenes de datos obtenidos de diversos modos. El análisis de datos no es nuevo ya se usaba la minería de datos. Lo que aporta la IA es la capacidad de extraer patrones que pueden servir para describir un determinado hecho en base al análisis de los datos, predecir comportamientos futuros basándose en comportamientos pasado o prescribir una determinada acción para conseguir un resultado.

Los grandes volúmenes de datos están asociados con la IA porque se necesitan para entrenar los algoritmos y, además, la eficacia y eficiencia de los algoritmos depende de la calidad de los datos sobre los que trabaje, de que hayan sido limpiados y preparados. Por ejemplo para aplicarlo en el reconocimiento facial y estado de ánimo del sujeto se necesita analizar millones de imágenes de rostros.

Servicio estatal de Promoción del Empleo (SEPE)

Ámbito: Empleo y Formación

Descripción del proyecto:

El objetivo del proyecto MDM es ofrecer definir e implantar una visión única de la información de la que dispone el SEPE del ciudadano, dando solución a problemas relacionados con la calidad del dato (registros duplicados, sin normalización, diferentes orígenes de datos) y aportando una visión completa de todos sus sistemas y la interacción entre ellos. Para ello se definirá un modelo de gobierno del dato adicionalmente a la implantación del conjunto de herramientas necesarias para la gestión de datos maestros.

Ministerio de Administración Pública de la República de Eslovenia

Ámbito: Empleo y Formación

Descripción del proyecto:

En 2017 Eslovenia, lanzó el proyecto piloto “Análisis de Big Data para la mejora de la eficiencia de los recursos humanos” Este proyecto piloto se lanzó para ver qué se podía hacer con las herramienta de Big data en cloud analizando los datos de recursos humanos del ministerio, para identificar potenciales mejoras de gestión de RRHH y posibles patrones de comportamiento.

Los resultados mostraron que existe una gran margen de mejora en el ámbito de la gestión de RRHH, por ejemplo en la planificación de las necesidades de contratación pública, con la ayuda de la IA.

Una de las herramientas más populares en la actualidad son las de análisis predictivo. Se utilizan, a menudo, como apoyo para tomar decisiones (decision-making) a partir de la fusión de datos en un dominio concreto aplicando algoritmos evolutivos. En este campo existen varias experiencias.

Intervención General del al Administración de Estado (IGAE)

Ámbito: Análisis de riesgos

Descripción del proyecto

Se va a evaluar la conveniencia de utilizar técnicas de IA (sistemas de apoyo a la decisión y análisis de sistemas dinámicos) como parte de un proyecto de análisis de riesgos para descubrimiento de fraude en subvenciones y análisis de riesgos para el Sistema de Supervisión Continua. Este proyecto se desarrollará durante la segunda mitad de 2019 y la primera de 2020.

Adicionalmente, se va a evaluar la posible utilidad de técnicas de procesamiento del lenguaje natural mediante un “chatbot” para el soporte a usuarios de aplicaciones. Este proyecto se desarrollará a finales de 2019 como un Trabajo de Fin de Master del curso selectivo del cuerpo TIC-A1, dirigido por personal de nuestra organización, y por tanto sin impacto presupuestario.

Machine learning

Es una rama de la inteligencia artificial cuyo objetivo es desarrollar técnicas que permitan que las computadoras aprendan, identifiquen patrones y tomen decisiones con una mínima intervención humana. Machine learning produce modelos de manera rápida y automática que puedan analizar grandes volúmenes de datos y producir resultados. El aprendizaje se realiza mejorando la experiencia de los modelos a base a iteraciones puesto que los modelos son expuestos a nuevos datos en cada una de ellas. Los algoritmos catalogados como Machine Learning se clasifican en dos grandes grupos: aprendizaje supervisado y aprendizaje no supervisado.

Ministerio de Industria, comercio y Turismo

Ámbito: Sector Turístico.

Descripción del proyecto:

TURESPAÑA cuenta con una base de datos totalmente anonimizada que aglutina las encuestas realizadas a turistas en los puntos fronterizos de España, a razón de unas 100.000 al año. Las encuestas son realizadas por el Instituto Nacional de Estadística (INE), bajo la denominación Encuesta de Gasto Turístico (EGATUR). Se pretende construir un prototipo que permita estudiar la viabilidad de clasificar a los turistas extranjeros en base a su comportamiento y a su perfil de gasto en España. Para ello, se emplearán técnicas machine learning de clasificación no supervisada, y en particular se usarán técnicas avanzadas de análisis clúster con el fin de obtener representaciones gráficas de los patrones de comportamiento y comprobar cómo evolucionan en el tiempo. Los patrones emergentes permitirán al decisor priorizar las políticas e iniciativas de promoción turística mediante técnicas de micromárketing dirigidas a los perfiles turísticos estratégicos.

Ministerio Público de la Ciudad de Buenos Aires

Ámbito: Justicia

Descripción del proyecto:

Prometea es el sistema de aprendizaje automático supervisado que ha utilizado la Corte Constitucional de Buenos Aires, que recibe miles de expedientes diarios, para reducir el tiempo que se gasta en seleccionar casos urgentes de 96 días a 2 horas. *Prometea* lee, analiza, detecta y en pocos segundos sugiere cuáles son los casos prioritarios en asuntos relacionados con la salud. Todo esto, bajo supervisión humana. La herramienta está diseñada para crear automáticamente múltiples documentos, pudiendo general automáticamente, 14 documentos en 16 minutos, cuando antes se tardaba 2 horas y 40 minutos una persona. El sistema combina procesamiento de lenguaje natural, automatización y la predicción.

Asistentes inteligentes chatbots

Un asistente personal inteligente es un agente tipo software que puede realizar tareas u ofrecer servicios a un individuo. Los más populares son los asistentes activados por voz como

Siri de Apple, Alexa de Amazon, Google assistant de Google y Cortana de Microsoft. Estos asistentes virtuales crean una interfaz conversacional mediante el uso de reconocimiento de voz e inteligencia artificial, y realizan tareas que, fueron diseñadas para hacerse a través de interfaces en donde el ingreso de información se realiza por medio de botones o entradas táctiles.

Un chatbot es una aplicación capaz de simular una conversación humana a través de una interfaz conversacional. Se desarrollan para una tarea específica y en general, los desarrolladores combinan técnicas de Machine Learning y de Procesamiento de Lenguaje Natural (NLP por sus siglas en inglés). Los chatbots han transformado la experiencia de los usuarios.

Consejería de Salud y Familia de la Junta de Andalucía

Ámbito: Sanidad, información a la ciudadanía

Descripción del proyecto:

Consiste en un asistente virtual para informar sobre el coronavirus, accesible a través de la App de Salud Responde y de la web de www. epes. es. Está basado en la tecnología de inteligencia artificial de IBM Watson, para ayudar a atender las consultas de los ciudadanos sobre la enfermedad del Covid-19.

Gobierno de Finlandia

Ámbito: Empresas

Descripción del proyecto:

“Starting up smoothly” es el chatbot puesto en marcha por el gobierno de Finlandia, para ayudar a los empresarios extranjeros. Les asiste para darles información de cómo establecerse y abrir sus negocios evitando los silos gubernamentales. Está compuesto por tres Chatbots: uno para ayudar a obtener permiso de residencia para vivir en Finlandia, uno para aconsejar cómo crear una empresa y otro para proporcionar información sobre impuestos laborales en el País.

Procesamiento de Lenguajes Naturales

El procesamiento del lenguaje natural, abreviado PLN —en inglés natural language processing, NLP— es un campo de las ciencias dentro de la inteligencia artificial y lingüística que estudia las interacciones entre las computadoras y el lenguaje humano. Se utiliza en los chatbot actualmente en diferentes áreas.

Agencia Estatal Tributaria

Ámbito: Fiscalidad y lucha contra el fraude

Descripción del proyecto:

Procesamiento de lenguaje natural: aplicación de chatbot para consultas de contribuyentes y extracción de información de fuentes no estructuradas para automatización de decisiones y lucha contra el fraude.

Machine learning: modelos predictivos para la automatización de decisiones y para la lucha contra el fraude en el ámbito tributario: devolución de IVA, alta en Registro de Operadores

Intracomunitarios, predicción de fraude en el ámbito del IVA, detección de relaciones anómalas entre contribuyentes que constituyen indicios de fraude.

Robótica: automatización de la decisión en múltiples procedimientos en el ámbito de la recaudación y la gestión de los tributos (actuación administrativa automatizada basada en reglas, si esto puede considerarse en esta categoría).

Sistemas de apoyo a la decisión: uso de modelos predictivos para el apoyo a la decisión de los actuarios en los ámbitos ya mencionados. Sistema de análisis de riesgos para la selección de contribuyentes.

Procesamiento de imágenes

Consiste en procesar señales, digitales, analógicas u ópticas y realizar con ellas operaciones matemáticas para tener como resultado una imagen mejorada que resulta más fácil de interpretar o reconocer. Es una herramienta ampliamente utilizada en muchas áreas, como la Visión Artificial o reconocimiento del entorno como sucede en el caso del vehículo autónomo para detectar obstáculos u otros vehículos.

Caso de uso del FEGA de procesamiento de imágenes

Ámbito: Agricultura y Ganadería

Descripción del proyecto:

Inspecciones previas al pago de ayudas basadas en Monitorización. Se trata de capturar imágenes del terreno mediante satélites Sentinel, que posteriormente son procesadas por sistemas de información que permiten discriminar si el uso que el agricultor ha declarado es conforme a la realidad. Para ello se utilizan técnicas de Aprendizaje automático (machine learning) que permiten ir evolucionando el reconocimiento del terreno en el tiempo.

Robótica inteligente (cognitiva)

La Robótica es la rama que se centra en la construcción de robots y aparatos que realizan operaciones o trabajos en sustitución de la mano de obra humana. El comportamiento de los robots se hace inteligente interactuando con el entorno vía sensores, actuadores y dotados de capacidad para la toma de decisiones, como sucede con robots empáticos antropomórficos en el acompañamiento de ancianos.

La irrupción de la IA en la robótica hace ya algunos años ha impulsado el desarrollo de «sistemas robóticos inteligentes» integrando las dos tecnologías. Esto ha permitido el desarrollo de robots autónomos (robots que pueden operar con un alto grado de autonomía que necesitan interpretar el contexto en el que se encuentran), que les permite acometer tareas complejas, capaces de percibir, razonar y actuar ante entornos dinámicos e imprevisibles o en entornos agresivos para el ser humano.

Ámbito: Agricultura

Descripción del proyecto:

El objetivo del proyecto ERMES es contribuir a mejorar la eficiencia de la agricultura y reducir las emisiones de pesticidas. En una industria donde gran parte del trabajo es repetitivo, intensivo y duro, la inteligencia artificial puede ayudar en varias áreas. Los robots inteligentes, como los del proyecto ERMES, pueden prever cuándo los cultivos están listos para la cosecha. Por otro lado, el cambio climático es otro de los campos en los que se puede aplicar la inteligencia artificial. Así, las supercomputadoras son capaces de combinar enormes cantidades de datos para crear modelos precisos del tiempo, además de predecir cómo serán los efectos del cambio climático a medio-largo plazo

Capacitación

El desarrollo de la inteligencia artificial requiere una sólida estrategia en aumentar las capacidades de los empleados públicos y del sector privado para hacer frente a la escasez de competencias. Un Plan de acción combinado y coordinado en inteligencia artificial de todas las unidades de formación de los Departamentos, el INAP y las Universidades contribuirá a identificar casos de uso de los datos en las tecnologías basadas en la inteligencia artificial.

INTEF

Ámbito: Formación

Descripción del proyecto:

Desde el área de Experimentación en el aula del INTEF se elaboran recursos que permiten tanto al profesorado como al alumnado convertirse en creadores de soluciones de inteligencia artificial, para que puedan conocer de un modo práctico y divertido el funcionamiento de este tipo de sistemas, en concreto los basados en aprendizaje automático (o machine learning), y ser conscientes tanto de las oportunidades como de los retos que plantea su uso en nuestro día a día.

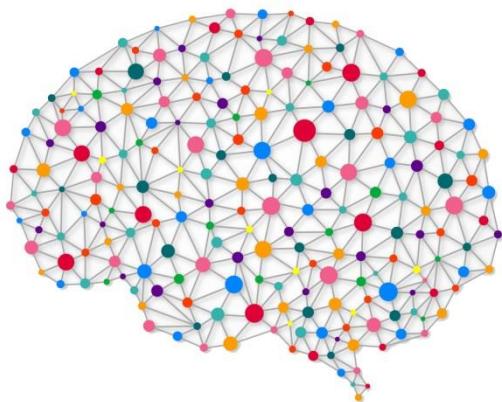
Así, para ilustrar el funcionamiento de los sistemas de inteligencia artificial basados en aprendizaje automático, se han elaborado un conjunto de vídeos en los que se desarrolla paso a paso un asistente virtual con el que controlar la iluminación y la temperatura de nuestro hogar usando órdenes en lenguaje natural.

Ver más: <http://code.intef.es/inteligencia-artificial-en-el-aula-con-scratch-3-0/>

Hay muchas otras áreas en las que la inteligencia artificial tiene aplicaciones, pero quizás una en la que más avances proyectan es el área de salud. La atención sanitaria se considera un sector que se beneficiaría enormemente de la IA. El mapa de capacidades de IA en España, registra un total de 98 entidades orientadas a la salud. En el caso concreto del sector sanitario, la tecnología actual permite algunas aplicaciones como sistemas de ayuda al diagnóstico clínico, indicando la probabilidad de diversas patologías en base a unas variables de entrada; sistemas de tratamiento de imágenes radiológicas (marcado automático de tumores sobre una imagen) y el diagnóstico automático de algunas patologías como, por

ejemplo, la retinopatía diabética basándose en imágenes de fondo de ojo. Estas aplicaciones son indudablemente útiles y poseen un gran beneficio potencial para los pacientes.

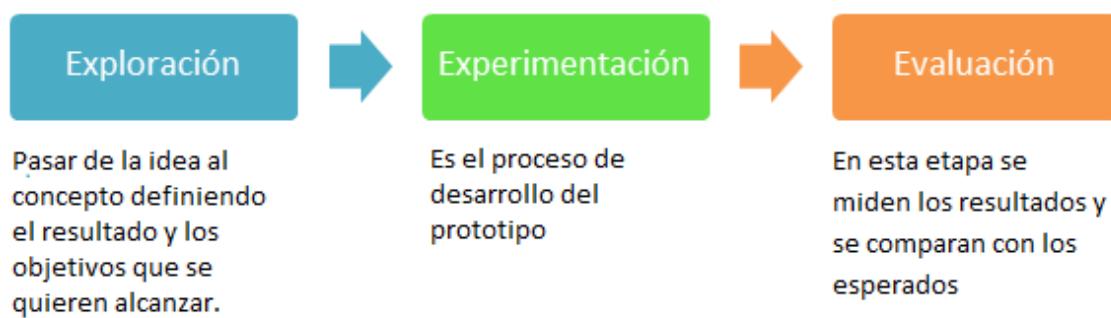
Metodología para aplicar la IA en un proyecto



desarrollo de productos innovadores.

A un proyecto de IA se le pueden aplicar metodologías y ciclos de desarrollo de software tradicionales. Nada lo impide. Pero teniendo en cuenta que un proyecto de IA es algo nuevo, probablemente está más cerca de un proyecto de innovación que de un proyecto clásico. Por esta razón esta guía describe una metodología específica para proyectos de innovación. Esta metodología propuesta toma su base de la metodología Stage-Gate® que proporciona un marco consistente y probado para gobernar el

En un proyecto de IA hay que definir, idear, prototipar, probar y redefinir. Así que la metodología propuesta estructura el proceso en **seis fases**, enmarcadas en 3 bloques que facilitan la construcción del sistema inteligente. Los tres bloques de construcción son los siguientes:



La metodología Stage-Gate® se plasma en un proceso de innovación de **seis etapas y cinco compuertas** (o puntos de decisión) y es apropiado para el desarrollo de nuevos productos y servicios en los que la tecnología, la fabricación y/o la comercialización es desconocida para la organización y, por lo tanto, requieren una mayor gestión de riesgos.

Otras metodología que se pueden utilizar son: el ciclo de vida descrito en el documento Artificial Intelligence in Society de la OCD; , más específicas de análisis predictivos, TDSP (el Proceso de ciencia de datos en equipo³) que define un ciclo de vida con los pasos completos, o puede utilizar otros ciclos de vida de la ciencia de los datos como Cross Industry Standard Process for Data Mining ([CRISP-DM](#)), Knowledge Discovery in Databases ([KDD](#)).

³ El proceso de ciencia de datos en equipo (TDSP) es una metodología de ciencia de datos ágil e iterativa para proporcionar soluciones de análisis predictivo y aplicaciones inteligentes de manera eficiente.

Factores para aplicar la IA en un proyecto

Un proyecto de IA se inicia para alcanzar múltiples objetivos como aumentar la productividad, dar mejor servicio o aumentar la calidad del servicio. En cualquier caso, para poder implementar en un proyecto de IA, se debe tener en cuenta **5** factores básicos y 4 más de reflexión. Los 5 factores básicos son:

1. **Tener un caso de uso en el que aplicar la IA.** A ser posible, comenzar con proyectos de poco alcance y buscar mejoras moderadas en los procesos, la satisfacción del usuario o la calidad.
2. **Tener datos y de calidad.** Para que un modelo de aprendizaje funcione se necesita tener un gran volumen de datos. A más datos más conexiones y mayor probabilidad de obtener precisión en los resultados o en patrones o en hacer recomendaciones. Sin embargo, tener datos sin depurar no va a facilitar estos resultados precisos. Por tanto hay que preparar esos datos. Contar con una política de datos ayudará en el proceso.
3. **Capacidad técnica.** Un proyecto de IA consume proceso y energía. Contar con una plataforma cloud para consumir, almacenar y analizar los datos va a ayudar.
4. **Conocimiento de la ciencia de los datos.** Será necesario saber gestionar, limpiar, integrar y transformar la información. Hay que estar capacitado para aplicar procedimientos estadísticos, reconocer datos relevantes y establecer indicadores.
5. **Contar con presupuesto.** Será necesaria una infraestructura de IA que habrá que construir, ejecutar y mantener y personal capacitado. El coste también va a depender de la estrategia para adquirir la tecnología, si se utilizarán APIs o algoritmos personalizados o desarrollarlos. Y lo mismo en cuanto a los sistemas, podría ser útil recurrir, inicialmente, a un sandbox en el que poder probar en un entorno sin necesidad de adquirir los sistemas de AI. También hay que tener en cuenta que los algoritmos, necesitan un tiempo de entrenamiento, supervisar los resultados, comprobar que no haya sesgos, etc.

Y los 4 factores a reflexionar:

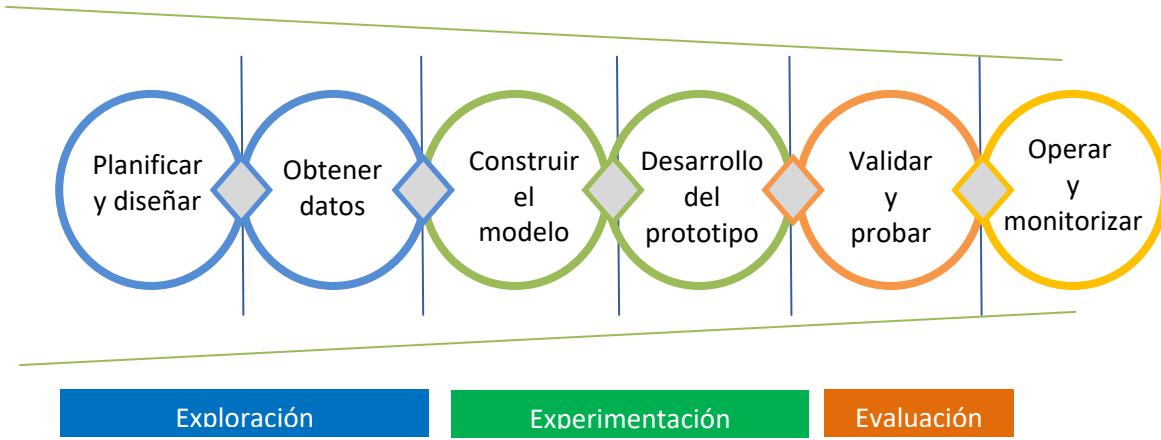
1. **Etica.** Es necesario plantearse la ética de usar los datos y que el modelo no sea perjudicial para ningún ser humano, ser vivo o al medio ambiente. Es decir, que cumpla con los principios éticos de la Unión europea.
2. **Privacidad.** Igualmente el modelo debe cumplir con el Reglamento de Protección de datos (GDPR) y la Ley orgánica de protección de datos.
3. **Responsabilidad.** hay que determinar quién es el responsable de los resultados y las decisiones del modelo.
4. **Explicabilidad y transparencia.** En la medida de los posibles y para generar confianza los desarrolladores y los responsables deben poder explicar cómo el modelo de IA ha llegado al resultado o al menos poder auditar el proceso.

Fases de la metodología en el proyecto IA

Para aplicar la IA en una organización antes ésta debe ser Data Driven.

No es recomendable empezar un proyecto de Inteligencia Artificial aplicando algoritmos a los datos sin antes determinar **los resultados y objetivos deseados**.

Se describen en este punto las **seis etapas y cinco compuertas** (o puntos de decisión). En cualquiera de las fases si el proyecto no convence puede tomar la decisión de retirar el sistema de IA.



0: Planificar y diseñar

Entender de la manera más completa el problema que se desea resolver, permitirá recolectar los datos correctos e interpretar correctamente los resultados. En la fase 0 lo primero es determinar qué caso de uso se puede llevar a cabo y definir los resultados que se quieren conseguir.

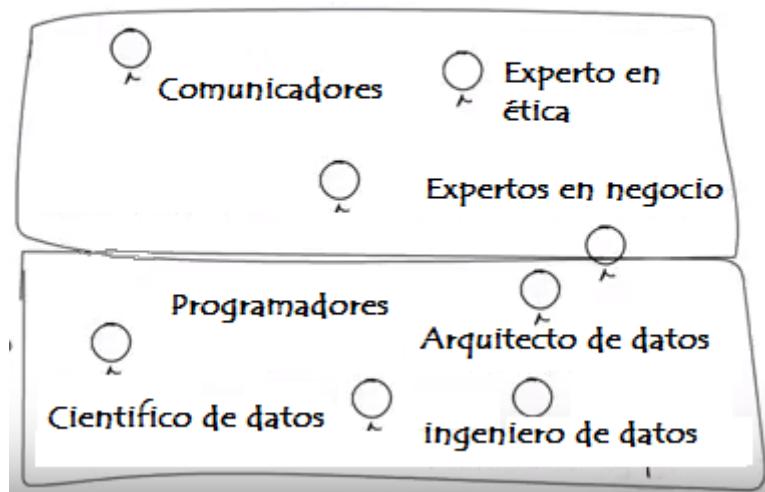
Comenzar con un proyecto con un alcance pequeño ayudará a aprender a utilizar la IA. Definir el proyecto buscando mejoras moderadas en los procesos, la satisfacción del usuario o la calidad. Una posibilidad es aplicar la IA a proyectos que ya se están usando un robot virtual o el análisis de datos para potenciar esas soluciones.

También, en esta fase hay que decidir distintos punto, entre ellos la estrategia de adquisición de los sistemas y herramientas:

- Construir o comprar las herramientas. Comenzar utilizando API de inteligencia artificial pre construidas necesita menos esfuerzo y conocimiento; puede ser un buen comienzo para introducirse en las soluciones de inteligencia artificial.
- Disponer de diversos sistemas como bases de datos, herramientas de análisis, algoritmos, plataformas de desarrollo, plataformas en cloud.

- Construir el equipo que va a necesitar el proyecto. Va a requerir contar con diferentes perfiles: personal con conocimiento de negocio, expertos en ciencia de los datos, expertos en informática, etc
 - **arquitecto de datos** para establecer la visión para el uso de datos por parte de la organización. Debe conocer las necesidades del negocio.
 - **científico de datos** para aprovechar el valor de los datos.
 - **ingeniero de datos** para desarrollar la entrega de productos de datos y servicios en sistemas y negocios procesos
 - **ingeniero de sistemas** con experiencia con contenedores y orquestaciones de contenedores - por ejemplo, Docker y Kubernetes
 - **Desarrolladores de software** con conocimientos de programación como Python, R u otro lenguaje de programación
 - **experto en ética** para proporcionar ética juicios y evaluaciones sobre las entradas del modelo AI
- Asignar responsabilidad de los modelos de IA
- Planificar el modelado de IA
- Definir los criterios de éxito.

EQUIPO



*Si el proyecto no le convence cierre en este punto y recopile la opinión de todos los involucrados.

1: Obtener los datos

En esta fase, se analizará el estado de los datos. Se procesará la recopilación y limpieza de datos, la realización de comprobaciones de integridad y calidad, y la documentación de los metadatos y las características del conjunto de datos. Se considerarán tres puntos: el volumen de los datos que debe ser suficiente como para aplicar IA; que sean datos relevantes; y que sean depurados para tener la calidad suficiente para obtener resultados válidos.

Conocimiento del negocio

El arquitecto de los datos se encargará de la preparación de datos que incluye las tareas generales de selección de datos a los que se va a aplicar una determinada técnica de modelado, la limpieza de datos, la generación de variables adicionales, la integración de diferentes orígenes de datos y los cambios de formato si fueran necesarios.

Será necesario definir las variables principales del modelo y las métricas asociadas para determinar el éxito del proyecto. Para ello se formulan las preguntas a los usuarios, los empleados y los involucrados.

- Cuánto? o ¿cuántos? (regresión)
- ¿Qué categoría? (clasificación)
- ¿Qué grupo? (agrupación en clústeres)
- ¿Es extraño? (detección de anomalías)
- ¿Qué opción se debe elegir? (recomendación)

Identificar las fuentes de los datos pertinentes a los que tiene acceso la organización o los que necesita obtener.

- ¿Hay indicadores para medir el objetivo y las características que están relacionados con él?

Hacer un catálogo de las fuentes de datos

Habrá que generar un conjunto de datos limpio y de alta calidad. Habrá de descubrir los patrones que son inherentes a los datos

Es esta fase conviene necesitará contar un experto con conocimiento del negocio, que conozca bien los datos y un experto en la ciencia de los datos⁴ (un data scientist). Estos profesionales tendrán conocimientos matemáticos y estadísticos, además de dominar el **software estadístico y la programación** y los sistemas de análisis de datos masivos. Su labor es convertir los datos en respuestas.

Al seleccionar los datos deberá asegurar que se cumple con los requisitos de protección de datos. La Agencia de protección de datos ha publicado Una aproximación para la adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial que se puede aplicar en este punto. Igualmente hay que asegurar que cumple con los principios éticos definidos por la Comisión Europea y que hace un uso responsable de los activos de los que dispone.

Puesto que el proceso va a ser iterativo se aconseja redactar un Informe de la calidad de los datos.

***Si no tiene datos cierre en este punto y recopile la opinión de todos los involucrados.**

⁴ Un científico de datos es sencillamente un profesional dedicado a **analizar e interpretar grandes bases de datos**

2: Construir el modelo. Fase de modelado

En términos matemáticos, se habla de “modelar” debido a que se crea un modelo, una reconstrucción simplificada de cómo funciona un proceso observado en el mundo real.

El modelado de datos puede ser utilizado para dos propósitos:

Predecir el valor de una variable resultante en base a valores conocidos. Estos modelos predictivos se emplean en una enorme variedad de aplicaciones: inversión en bolsa, prevención de fraude, publicidad online, fijación de primas en seguros de riesgo, etc.

Explicar la relación entre una variable dependiente y todas las demás (las explicativas), buscando determinar si la relación es significativa. Los modelos explicativos son los que se usan en investigación académica, ya que ayudan a entender el fenómeno modelado.

Se debe dividir los datos de entrada aleatoriamente y seleccionar la muestra que se va a utilizar para el entrenamiento y la correspondiente para la prueba.

3: Desarrollo

* Este es un proceso que hay que repetir si no se obtiene el resultado.

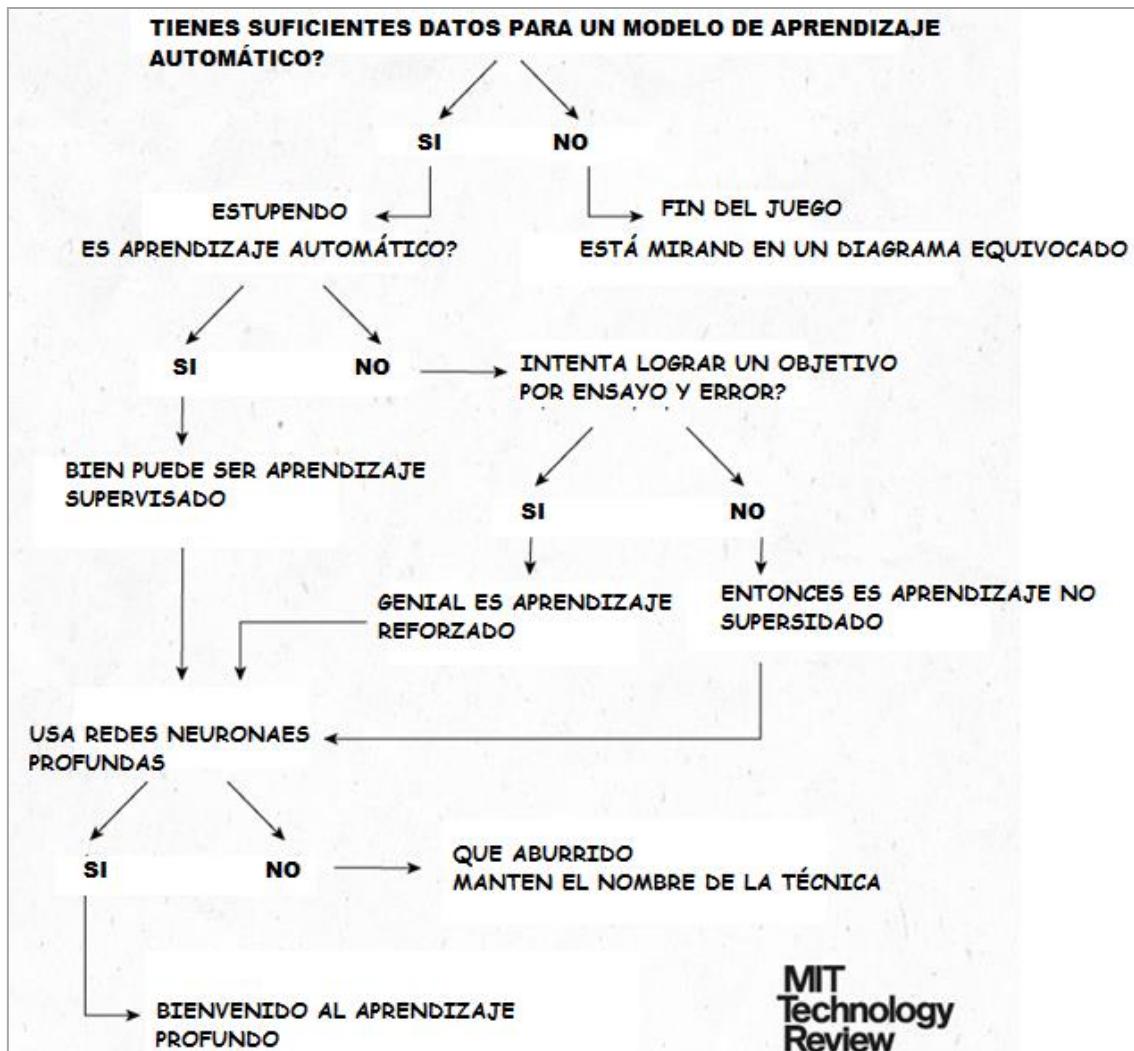
El objetivo de esta fase es conseguir un prototipo, que será ampliamente probado durante la próxima etapa. Para cada modelo que se prueba, conviene generar un informe del modelo detallado sobre cada experimento.

Si los datos de entrenamiento están sesgados, el modelo resultante también lo estará.

Según el tipo de la pregunta que intenta responder escogerá entre los múltiples algoritmos de modelado disponibles. Es en esta fase en la que necesitará un experto informático y decidir si utilizará algoritmos pre construido o no.

Entrenar el modelo con el conjunto de datos de *entrenamiento* hasta que responda a la pregunta con la máxima precisión comparando con las métricas de éxito. Este entrenamiento consiste en ir alimentando a los algoritmos con un **conjunto de datos “conocidos”**, hasta que se obtengan unos resultados adecuados. El resultado de este proceso será un **modelo**, es decir, **conjunto de algoritmos que han sido entrenados con datos**.

El siguiente esquema nos ayudará a analizar si tenemos datos para un proyecto de IA.



*La puerta de decisión se abrirá cuando el producto se haya desarrollado lo suficiente.

4: Prueba y validación

Probar el modelo. Para ello hay que pasar el conjunto de datos *de pruebas*,

Esta fase abarca las pruebas y la validación de productos. Durante este período hay que completar una serie de subetapas:

- **Identificar errores.** El propósito de esta prueba es identificar posibles errores de producción u otros problemas.
- **Pruebas con la ayuda de usuarios.** En esta parte, el artículo es probado por varios participantes.

Comprobar que los resultados obtenidos son satisfactorios. Examinar los indicadores de éxito del sistema previamente definidos.

*Si las pruebas no dan buenos resultados, la puerta de la siguiente fase permanecerá cerrada y deberá hacer otra iteración.

5: Lanzamiento e implementación

Un modelo que funciona bien puede ser consumido por otras aplicaciones. En esta fase hay que comprobar la compatibilidad con los sistemas heredados, garantizar el cumplimiento normativo, gestionar los cambios organizativos y evaluar la experiencia del usuario.

Si el producto está listo para ser lanzado, debe abordar la puesta en producción y la comunicación. Por tanto, en esta fase hay que calcular el coste de mantener el modelo.

Durante la fase de funcionamiento y operación ha de aplicar una evaluación continua. Es necesario observar las consideraciones éticas comprobando que no tienen sesgos y coincide suficientemente con las necesidades y los deseos del usuario. El modelo deberá funcionar con datos “desconocidos”.

* En esta fase se identifican los problemas y se reajusta el modelo, volviendo a las fases anteriores, si fuese necesario o retirando el modelo de IA de la producción.

Finalmente, este primer modelo le servirá para crear otros más complejos. Es probable que con el tiempo necesite un reentrenamiento.

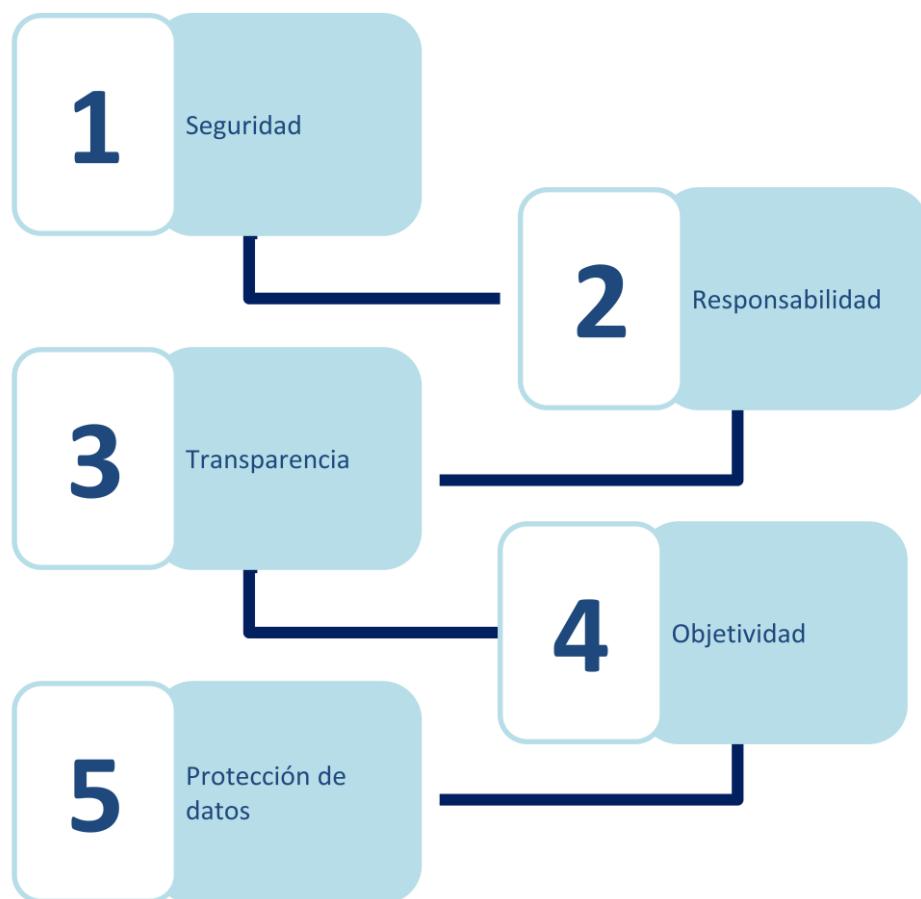
En el bloque siguiente se ofrece una lista de que ayuda a revisar el impacto que su proyecto puede tener sobre determinados aspectos importantes.

Impacto.

Es difícil establecer el impacto de la IA en la administración pública habida cuenta de la gran velocidad a la que avanzan este tipo de sistemas. No obstante, cualquier sistema informático de que se dote la administración pública para apoyarse en la toma de decisiones deberá salvaguardar al menos los mismos valores y estándares en los que hoy se apoya, tales como la transparencia, la responsabilidad, la no discriminación, la protección de datos y la seguridad, entre otros.

Así, en relación a estos sistemas de IA utilizados por las administraciones públicas deberá ser posible garantizar el cumplimiento de leyes, normas y valores ya existentes en diversos ámbitos como los siguientes.

IMPACTO DE LA IA



Seguridad

Del mismo modo que en los sistemas informáticos que ya se vienen utilizando en las administraciones públicas, en los sistemas apoyados en IA deberán realizarse todas aquellas acciones encaminadas a garantizar la seguridad de los mismos, en particular las siguientes:

- **Evaluación de riesgos** para la administración pública en el diseño y desarrollo de sistemas de IA.
- **Evaluación continua** de que los sistemas de IA funcionan de manera adecuada a su finalidad prevista.
- Establecimiento de **mecanismos de supervisión** que permitan que los sistemas de IA **puedan ser examinados y evaluados adecuadamente**.
- **Adquisición de sistemas informáticos externos.** Requisitos de adquisición que aseguren que las empresas que desarrollan estos sistemas para el sector público cumplen adecuadamente los estándares públicos. Para ello se añadirán disposiciones y cláusulas en los pliegos de contratación atendiendo especialmente a la protección de colectivos diversos así como evitar el establecimiento de brechas digitales, discriminación y parcialidad.

Responsabilidad

La propia naturaleza de la IA puede hacer difícil establecer sobre quién recae la responsabilidad de las decisiones tomadas en base a estos sistemas. A pesar de ello habrá de ser posible identificar la atribución de responsabilidad de los organismos y funcionarios públicos. De este modo, deberá asignarse y documentarse en las etapas de diseño dónde recae la responsabilidad de los sistemas de IA de forma que las personas operadoras de los sistemas de IA puedan ejercer su función de forma efectiva.

Si un sistema de inteligencia artificial toma decisiones, deberá asegurarse que puede explicar los parámetros que han llevado a la toma de decisiones en el ámbito de los procesos administrativos y quien es el responsable de esa decisión. Del mismo modo que ocurre en la automatización de los procesos burocráticos de las administraciones públicas a través de la actuación administrativa automatizada, en el caso de la IA siempre deberá ser posible establecer previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente; así como señalar el órgano que debe ser considerado responsable a efectos de impugnación (art. 41 Ley 40/2015).

Por su parte las administraciones públicas deben asegurar que los funcionarios responsables de sistemas de IA conocen esos sistemas y reciben una formación continua.

Esa responsabilidad comienza en las etapas de diseño del sistema, en la validación y las pruebas. En muchos casos un humano supervisará las decisiones.

Ética

Según la UE, **cualquier AI debe respetar la autonomía y derechos de los ciudadanos europeos**, prevenir cualquier tipo de discriminación y ser imparcial.

La UE tiene el objetivo de promover una inteligencia artificial fiable. La IA no es magia y hay que asumir la responsabilidad de crearla, aplicarla y regularla de manera ética.

Cuando se habla de ética siempre se piensa en los problemas del sesgo, pero no se tienen en cuenta el consumo de energía desbocado que el entrenamiento de grandes volúmenes de datos representa y que por tanto hay que plantearse la sostenibilidad de algunas tecnologías de IA. Otro aspecto es el mal uso que se le puede dar.

En el anexo se incluye un cuestionario para construir una IA fiable.

Transparencia

En los sistemas de IA, que manejan ingentes volúmenes de datos y en otros tales como aquellos que evolucionan y aprenden, puede ser extraordinariamente difícil conocer cuál ha sido el proceso de toma de decisión en los mismos. No obstante, para salvaguardar aspectos éticos, de transparencia y de responsabilidad, cualquier sistema que tome decisiones apoyado en la IA deberá poder proporcionar algún sistema de explicación para informar de cuál ha sido el proceso que le ha llevado a tomar una decisión.

En aquellos sistemas que hagan muy difícil o imposible la posibilidad de encontrar una explicación a este proceso de toma de decisión, la administración deberá justificar de forma legítima la necesidad de utilizar este tipo de sistemas inexplicables que no cumplen el requisito de transparencia.

En todo caso siempre deberá ser posible ofrecer explicaciones en áreas tales como la sanidad, la seguridad y aquellas que se refieran a derechos del ciudadano que justifiquen con claridad esas decisiones dando cuenta de los fundamentos, razones, circunstancias concretas de una decisión automatizada específica.

Es probable que el análisis previo de los sistemas basados en IA en las etapas de diseño para poder dar cumplimiento a este objetivo de transparencia, es probable que también repercuta en un aumento de la transparencia en entornos donde no se utilice la IA.

Objetividad.

Ya hemos referido que una de las ventajas de la IA es poder trabajar con grandes repositorios de datos. Uno de los riesgos asociados a los resultados de estos análisis es que aunque los volúmenes de información sean enormes, los datos estén sesgados, por lo que las conclusiones extraídas de los mismos también lo estarán. El mismo problema aparece cuando se producen errores de muestreo en los datos que se utilizan para entrenar sistemas basados en IA, lo cual producirá también resultados sesgados debido a ese error.

Otro riesgo de sesgo es el que se incorpora en el diseño de los algoritmos, ya que los humanos, por el hecho de serlo, incorporamos consideraciones individuales o de grupo, valoraciones, políticas, prejuicios, intereses, éticas o juicios de todo tipo. Por este motivo la IA puede convertirse en un amplificador de sesgos y desigualdades ya presentes en aquellos que han diseñado los sistemas, por ejemplo en aquellos sistemas de IA que aprenden en base a estas reglas ya sesgadas.

Cabe preguntarse si es realmente posible que a partir de los datos disponibles o a partir de seres humanos sea posible crear sistemas no sesgados, pero, en todo caso, es importante tener en cuenta estas consideraciones para al menos minimizar o tener en cuenta el impacto del sesgo en los sistemas basados en IA, de cara a enfocarse más en su eficacia que en su infalibilidad; esto es, hay que asumir que en estos sistemas habrá un cierto sesgo en los resultados obtenidos.

Algunas recomendaciones para minimizar el sesgo son por ejemplo utilizar equipos humanos en el diseño lo más diverso posible, para trasladar esta diversidad a los algoritmos construidos; también realizar estudios y análisis para encontrar sesgos en los resultados para en su caso corregir los algoritmos; también analizar si las muestras y los orígenes de datos son suficientemente confiables en términos de sesgo de las variables objeto de trabajo.

Protección de datos y regulación.

Probablemente no sea necesaria una nueva regulación en esta materia, aunque sí que habrá que estudiar la adaptación de parte de la regulación vigente en áreas tales como la transparencia y el sesgo de los datos.

Por su parte, los proveedores de servicios (públicos y privados) deberán informar a los ciudadanos de que están recibiendo un servicio asistido por una IA y así como el método de recurso contra decisiones automatizadas asistidas por la IA.

La Agencia de protección de datos ha publicado la guía: *“una aproximación para la adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial”* que los técnicos de datos deben conocer.

Anexo 1 Glosario inicial de la IA

Inteligencia artificial

Es un campo de las ciencias de la computación que investiga y desarrolla sistemas capaces de realizar tareas inteligentes. El término fue acuñado por el informático y matemático estadounidense John Mc Carthy en la seminal Conferencia de Dartmouth de 1956. La propuesta para la conferencia era:

“proceder sobre la base de la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, ser descrito de forma tan precisa que se puede lograr que una máquina la simule”.⁵ Los participantes de la conferencia no acordaron sobre una teoría general en el campo pero sí compartían la visión de que las computadoras podían ser programadas para realizar tareas de modo inteligente.

En las décadas siguientes, el campo de estudio se expandió y desarrolló de la mano de teóricos tales como Alan Turing y Marvin Minsky, con diversos períodos de expansión y retracción que culminaron en el boom del siglo XXI producto del mayor poder computacional y cantidad de datos disponible.

Machine Learning (Aprendizaje automático)

La IA es diferente del aprendizaje automático o el aprendizaje profundo. Machine learning representa una de algunas de las técnicas y herramientas utilizadas en IA.

Es rama de la IA que tiene como objetivo hacer que los sistemas mejoren su desempeño (o “aprendan”) para realizar una tarea o resuelvan un problema de manera automática, o sea, sin una programación explícita, y en base a una gran cantidad de datos y según una medida de evaluación. El aprendizaje puede ser: supervisado (deducir una función a partir de datos de entrenamiento con inputs y outputs), no supervisado (construir una función a partir de datos de entrada) o por refuerzo (entrenamiento para tomar determinado tipo de decisiones). El término fue creado por el estadounidense Arthur Lee Samuel en 1959.

Deep Learning (Aprendizaje profundo)

Es un tipo de machine learning que utiliza una red neuronal artificial que imita el funcionamiento del sistema nervioso humano. La red tiene niveles jerárquicos que envían información al nivel siguiente para una elaboración más compleja.

Se utiliza para identificar características o patrones en objetos o datos no estructurados. La aplicación más conocida es el reconocimiento de imágenes y audio.

Transfer learning

Es un tipo de machine learning en el que un modelo desarrollado para una determinada tarea se usa como punto de partida para otra tarea o problema.

Algoritmo

Un proceso o conjunto definido de reglas que permiten llevar a cabo una operación o resolver un problema. Dado un estado inicial, se transita por una serie de estados sucesivos hasta llegar a un estado final. Se utiliza para el procesamiento de datos, cálculos y problemas matemáticos o computacionales. Es un elemento clave de la IA.

Big Data

Se utiliza el término Big Data para referirse a un conjunto de datos (estructurados o no estructurados) de tamaño y complejidad tal que no puede ser procesado por las herramientas de software tradicionales, y que tiene el potencial de ser analizado para la extracción de información útil. Según IBM, el Big Data se caracteriza por las “3V”: Volumen, Velocidad y Variedad, a las que luego se agregan otras dos, Valor y Veracidad.

Minería de datos

Es un campo de las ciencias de la computación que busca descubrir patrones y tendencias en grandes volúmenes de datos

Referencias:

Communication: Building Trust in Human Centric Artificial Intelligence. (2019, 8 abril). UE.
<https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence>.

LIBRO BLANCO sobre la inteligencia artificial. (2020, 19 febrero). UE.
https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf

<https://docs.microsoft.com/es-es/azure/machine-learning/team-data-science-process/lifecycle>

Tutoriales: <https://docs.microsoft.com/es-es/azure/machine-learning/team-data-science-process/walkthroughs>

Metodología para el desarrollo de proyectos en minería de Datos CRISP-DM. Basado en a tesis de José Alberto Gallardo

GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL
CREADO POR LA COMISIÓN EUROPEA EN JUNIO DE 2018

The Living Lab Methodology Handbook. Anna Ståhlbröst and Marita Holst, Social Informatics at Luleå University of Technology and CDT – Centre for Distance-spanning Technology, Sweden.

A guide to using artificial intelligence in the public sector. Office for artificial intelligent. Government Digital Service.

Artificial Intelligence and Public Standards. A Review by the Committee on Standards in Public Life. Chair, Lord Evans of Weardale KCB DL

Artificial intelligence and its use in the public sector. OECD.

Artificial Intelligence for Citizen Services and Government Harvard Ash Center Technology & Democracy Fellow